Improving WorkflowEfficiency for Mammography Using Machine Learning

Trent Kyono[a,1], Fiona J. Gilbert[b,c], Mihaela van der Schaar[a]

[a] *Department of Computer Science, University of California Los Angeles,*

291 Engineering VI, Los Angeles, CA 90095-1596, USA

[b] *Department of Radiology, University of Cambridge School of Clinical Medicine, Box 218, Cambridge Biomedical Campus, Cambridge, CB2 0QQ, UK*

[c] *NIHR Cambridge Biomedical Research Center, Box 277, Hills Road, Cambridge, CB2 0QQ, UK*

We know of no conflicts of interest associated with this publication, and there has been no significant financial support for this work that could have influenced its outcome.

Abstract

Objective: The aim of this study was to determine whether machine learning could reduce the number of examinations the radiologist must read by using a machine learning classifier for interpreting negative mammograms that it is confident in, and offloading or deferring the uncertain predictions to a radiologist for interpretation.

Methods: Mammograms were obtained from a private imaging dataset containing over 7,000 patients collected through six NHS Breast Screening Program centers throughout the UK. A convolutional neural network in conjunction with multi-task learning was used to extract imaging features from mammograms that mimic the radiological assessment provided by a radiologist, the patient's non-imaging features and pathology outcomes. A deep neural network was then used to concatenate and fuse multiple mammogram views to predict both a diagnosis and a recommendation of whether or not additional radiological assessment was needed.

Results: 10-fold cross-validation was used on 2000 randomly selected patients from the dataset; the remainder of the dataset was used for

convolutional neural network training. While maintaining an admissible NPV of 0.99, the proposed model was able to reduce 34% (95% confidence interval, 25-43%) and 91% (95% CI: 88% - 94%) of the negative mammograms from the radiologist for test sets with a cancer prevalence of 15% and 1%, respectively.

Conclusion: Machine learning was leveraged to successfully reduce the number of negative mammography examinations that radiologists need to examine without degrading diagnostic accuracy.

Introduction

The total number of screening mammography examinations conducted in the US alone is nearly 40 million annually and increasing [1]. Moreover, as examination volumes and time of interpretation increase with newer screening technologies such as digital breast tomosynthesis (DBT), radiologists will be forced to read more patients in less time [2]. Since a large majority of mammograms a radiologist examines are negative, machine learning methods that could triage a subset of examinations as negative with extremely high accuracy and refer the rest to a breast imager could significantly reduce the daily

interpretive workload of radiologists, freeing up time to focus on more suspicious examinations and diagnostic work-ups.

Recent advances in artificial intelligence, particularly deep learning, have led to significant improvements in computer-aided diagnosis (CAD) and decision support [3] but have not yet provided mechanisms for effectively reducing the number of examinations that a radiologist reads. The most popular CAD-based methods are centered around improving detection and diagnostic performance through CAD [4], but still require manual examination and validation by an expert radiologist. This comes with the overhead of additional CAD software interfacing for the radiologist, which has been shown to significantly increase the average reading time per patient [5].

Unlike related works using convolutional neural networks (CNNS) for mammography that attempt to completely automate and replace the radiologist [6-12], we depart from these methods by exploring a more conservative approach. Specifically, we explore a hybridized approach of mammography triage where some mammograms are autonomously diagnosed as being negative by a machine learning classifier and the remaining exams are read by a radiologist.

To address this, we extended and modified the implementation of Man and Machine Mammography Oracle (MAMMO) that was presented in Kyono et al. [13]. MAMMO was originally developed as a clinical decision support system that aimed to reduce the number of patients (both positive and negative) the radiologist read by relying on the decisions of a machine learning classifier for diagnosing mammograms that it was confident in, and offloading or deferring the uncertain decisions to a radiologist. However, in this study we redesigned MAMMO into a new system (Figure 1), called Autonomous Radiologist Assistant (AURA), which aimed to reduce the negative patient workload for the radiologist while sustaining a high negative predictive value (NPV).

The purpose of this proof-of-concept study was to determine whether the AURA system could autonomously triage patients between a machine learning classifier and a radiologist to effectively reduce the number of negative mammograms that a radiologist reads while maintaining an admissible NPV (greater than 99%).

Methods

Study Population

The *Tommy* dataset was originally compiled to determine the efficacy and diagnostic performance of DBT in comparison to digital mammography and was collected through six NHS Breast Screening Program (NHSBSP) centers throughout the UK [14]. It is a rich and well-labeled dataset with over 7,000 patients (over a 1,000 malignant) who received diagnostic mammograms, and includes radiological assessments, density estimates (μ = 38.2, σ = 20.7), age at examination (μ = 56.5, σ = 8.75), and pathology outcomes from core biopsy or surgical excision. The *Tommy* dataset does not include ethnicity or socioeconomic breakdown, and was therefore not factored in this study.

Although not all patients in the *Tommy* dataset underwent biopsy or had a later follow up examination, each patient underwent expert radiologist interpretation of both DBT and mammography modalities that significantly reduced the likelihood of false negative readings by

as much as 15-30% [14]. Patient distributions for age, breast density, and dominant radiological features are shown in Table 1.

The *Tommy* dataset was designed to challenge the radiologist with overlapping breast tissue cases. The patient criteria for selection were one of the following: 1) women recalled after routine breast screening between the ages of 47 and 73, or 2) women with a family history of breast cancer attending annual screening between ages of 40 and 49. Mammograms were read by over thirty radiologists with at least 2 years of experience reading 5000 mammograms or more per year.

Multi-task feature extraction

The ability for CNNs to learn complex spatial relationships and learn subtle and intricate pixel-based patterns make them a perfect tool for learning from radiological images [15]. The first step involved training a CNN using multi-task learning (MTL) to predict both the diagnosis and radiological assessments given a single mammogram image. Syeda-Mahmood [16] demonstrated the importance of incorporating clinical knowledge with medical imaging to improve clinical inference. Motivated by this, MTL was used to predict the radiological assessments for each mammogram to learn refined

feature representations and improve classification performance of the primary task, diagnosis, by obligating the CNN to learn the radiological assessment known to be associated with cancer. There were five other auxiliary tasks, which are shown as the outputs of Step 1 (Figure 1) and include the mammographic sign (e.g. circumscribed mass, spiculated mass, asymmetrical density, etc.), mammographic suspicion (similar to BI-RADS), mammographic conspicuity (either not visible, barely visible, not clear, or clearly visible), estimated breast density from a 10-cm VAS (visual analogue scale), and patient age at reading. Both breast density and patient age were included as auxiliary tasks because of their known associations with breast cancer [17, 18]. Training details regarding CNN architecture, hyperparameters, and image preprocessing were implemented from Kyono et al. [13].

Autonomous diagnosis

The second step of AURA involved autonomous diagnosis of a patient by considering all four mammogram views, rather than one view as done in Step 1. This was done by training a classifier that takes as input the CNN predicted multi-task outputs (MTO) for each of a

patient's mammogram views and their non-imaging features to issue a patient-level diagnosis. We chose to combine multiple mammogram views by concatenating over the CNN predicted MTO, rather than the penultimate dense layers as done by Carnerio et al. and Geras et al. [8,19]. This provided a couple noteworthy advantages. First, the MTO were extracted imaging features that emulated the radiological assessment and are what the radiologist would naturally consider when reading multiple mammogram views, such as asymmetries in breast density between left and right mammogram views [20]. Second, by pre-training our CNN in the first training phase, the MTO served as a refined feature space for combining mammogram views that improves performance in limited data scenarios commonly encountered in the medical imaging domain.

Autonomous assessment

The objective of Step 3 was to determine whether the diagnostic prediction made in Step 2 could be trusted, while also considering the patient's non-imaging features, such as age, and the radiological predictions of the CNN from Step 1. The primary design goal of AURA was to reduce the number of negative mammograms the radiologist

interprets by offloading examinations to a machine learning classifier (Step 2) while preserving a desirable NPV (greater than 99%). It is important to note that the NPV of AURA is representative over the collaborative system comprised of AURA autonomously diagnosing some partition of the cases and the radiologist diagnosing the rest. A loss function presented in Kyono et al. [13] was used to satisfy this constrained optimization that takes into consideration the performance of the radiologist. The uncertainty of decisions was calculated using dropout and test-time augmentation [21, 22]. Dropout was performed at a rate of 0.4 per dense layer and test-time augmentation was performed at 50 random samples per patient.

Validation of Performance

To evaluate the performance of AURA for *patient triage*, experiments were conducted on the *Tommy* dataset. 10-fold cross-validation was performed on 2000 randomly selected patients from the *Tommy* dataset and the remainder was used for CNN training. Specifically, the dataset was split into 10 even groups (folds), with each model tested on one fold, validated on another, and trained

on the remaining folds. This was repeated 10 times over each unique test set to provide a more stable and bias free model estimate. In deep learning, a validation set is used for determining model fitness and to find a stopping point for model training, while a test set is reserved for evaluating model performance.

To ensure the 2,000 patient subset was a proportional subset of the entire dataset, we maintained the dataset's 15% malignancy rate through the test set as well, by randomly sampling 300 malignant patients and 1700 normal patients from the entire dataset. Each of the cross-validation folds maintained the same 15% cancer prevalence as well.

To demonstrate the impact of cancer prevalence on the outcome of patient triage by AURA, AURA was applied on reduced cancer prevalence subsets, where cancer patients were randomly removed from each cross-validation test set to obtain the desired prevalence, specifically 1% and 5%.

Models were saved and evaluated at the point with the highest area under the receiver operating characteristic curve (AUROC). The performance of AURA *patient triage* was evaluated by investigating the number of patients that AURA deferred to the radiologist while

still satisfying an admissible NPV (greater than 99%). Additionally, to better understand the patient distributions that AURA successfully diagnosed, an evaluation of patient triage by attributes, such as breast density or age, was investigated. The odds ratio for each patient demographic examined was calculated using logistic regression analysis to identify the likelihood of AURA correctly identifying negative cases.

Results

For mammograms to autonomously bypass radiologist viewing, AURA must perform at some minimum NPV. In clinical practice, a desirable NPV is greater than 99%. Figure 2 shows various desired NPV operating points with the respective maximum number of patients that AURA was able to filter from the radiologist while maintaining the desired NPV (*x*-axis) for three stratifications of cancer prevalence (15%, 5%, and 1%). It is important to note that the reported NPV of the AURA system is comprised of both the AURA machine learning classifier and a radiologist.

At a cancer prevalence of 15%, AURA is expected to correctly diagnose approximately 34% (95% CI: 25% - 43%) of the patients while maintaining a NPV of at least 99%. The number of patients AURA filters are significantly improved when considering a cancer prevalence closer to screening distributions such as the 1% prevalence (Figure 2), where nearly 91% (95% CI: 88% - 94%) of the patients are screened from the radiologist. In summary, the data in Figure 2 shows that the number of patients that AURA filters is increased by either lowering the NPV threshold or by a decrease in cancer prevalence. Since the NPV threshold is likely not to change, this suggests that AURA should be favored in populations with lower cancer prevalence, i.e., screening populations.

AURA had a higher odds ratio of negative classification and NPV for patients that had attributes associated with lower likelihoods (or suspicion) of cancer, i.e., patients with lower breast density, patients with low visual conspicuity in their mammograms (suspicious regions are not clearly defined), patients with no suspicious imaging features in their mammograms, patients recalled by assessment (rather than family history), and younger patients (Table 2).

Similar to the odds ratio and NPV results, when investigating the patient populations that AURA filtered (and was confident in diagnosing) from the radiologist (Table 2), it was discovered to correlate with the same attributes associated with lower likelihoods of cancer, except for breast density. AURA chose to autonomously diagnose a higher percentage of patients with breast densities in the 25% to 50% range. This is justified by a lower cancer prevalence and a higher demographic of these cases existing in the dataset population (Table 1). The latter would result in more samples available for training and improving the AURA classifier.

Discussion

In this work, we showed the potential for improving the radiologist workflow using machine learning that is currently not explored in the current literature and practice. We presented the AURA system on a private mammography dataset and demonstrated that the AURA system could autonomously triage patients between a machine learning classifier and a radiologist to effectively reduce the number of negative mammograms that a radiologist reads while maintaining an

admissible NPV (above 99%). It was discovered that AURA filtered patients from the radiologist with attributes known to be associated with lower likelihoods of cancer, such as younger age, lower breast density, etc. (Table 2).

As far as we know, this is the first study investigating machine learning methods for reducing the number of negative mammograms a radiologist needs to read without replacing the radiologist entirely. In terms of neural network architecture, the closest work to ours is the prior study by Kooi and Karssemeijer [9], which investigated the impact of dataset and image size on predicting BI-RADS.

AURA may provide several benefits for radiologists. First, we found that AURA was capable of significantly decreasing the workload for the radiologist by 34% (95% CI: 25-43%), in a diagnostic setting (15% cancer prevalence) and by nearly 91% (95% CI: 88% - 94%) for a screening-like setting (1% prevalence). Although the *Tommy* dataset is comprised of diagnostic mammography and a direct comparison cannot be drawn to screening mammography directly, the 1% cancer prevalence results show significant promise for applications to screening mammography. Second, AURA filters the patients with lower likelihoods of cancer from the radiologist, leaving them with

more time to focus on the difficult cases that AURA is uncertain about. This additional time will allow for better scrutinization of the difficult cases, and may reduce the likelihood of misdiagnoses [23]. Lastly, AURA provides a palatable solution for incorporating artificial intelligence into radiology that does not completely remove the clinician from a radiological workflow and strives to do so for only particular examinations.

The current study has several limitations. First, AURA experiments were conducted on a mammography dataset originally collected for a reader study comparing DBT and digital mammography performance and is comprised of a high concentration of patients with overlapping breast tissue. Because of this, it is expected that AURA performance would improve significantly if performed on a screening population. Second, AURA would have benefited from increased image quantity and size (for CNN training), which has been demonstrated by Kooi and Karssemeijer [9] to improve classification performance in deep learning for mammography.

Lastly, our dataset did not contain annotated regions of interest (ROI), and was therefore not leveraged in AURA's training regime and architecture. ROI-based machine learning methods have been shown

to improve performance on mammograms on several small datasets [6-12, 24-30], and would be expected to improve classification performance of AURA as well.

In addition to validation on larger, population-based imaging datasets, future research directions include investigating AURA's potential for patient triage based on differing levels of suspicion for cancer. Such a triage system would allow examinations that are most likely to have cancer to be interpreted earlier by the radiologist and expedite faster recall and diagnostic evaluation [31]. Similarly, patients with a higher likelihood of cancer could be seen earlier in the day when the radiologist is more alert, thus mitigating misdiagnoses associated with fatigue [23]. Lastly, due to the visual similarity between DBT images and mammograms, AURA can easily be extended and tested on DBT.

In conclusion, our study demonstrates a proof-of-concept for machine learning to be used for improving breast imaging workflow for mammogram interpretation. AURA opens the door for realistic synergistic relationships between radiologist and machine with benefits that surpass those reported in the existing literature, and

provides methods for artificial intelligence integration that can be integrated into clinical practice in the near term.

References

[1] Broeders M, Moss S, Nystrm L, et al. The impact of mammographic screening on breast cancer mortality in europe: A review of observational studies. *Journal of Medical Screening* 2012;19:14–25.

[2] Sokolovskaya E, Shinde T, Ruchman RB, et al. The effect of faster reporting speed for imaging studies on the number of misses and interpretation errors: A pilot study. *Journal of the American College of Radiology* 2015;12(7):683 – 688.

[3] Giger ML. Machine learning in medical imaging. *Journal of the American College of Radiology* 2018;15(3, Part B):512 – 520.

[4] Rodriguez-Ruiz A, Mordang JJ, Karssemeijer N, Sechopoulos I, and Mann RM. Can radiologists improve their breast cancer detection in mammography when using a deep learning based computer

system as decision support? *Proc.SPIE* 2018;10718:10718 – 10718 – 10.

[5] Tchou PM, Haygood TM, Atkinson EN, et al. (2010). Interpretation time of computer-aided detection at screening mammography. *Radiology* 2010;257(1):40–46.

[6] Akselrod-Ballin A, Karlinsky L, Hazan A, et al. Deep learning for automatic detection of abnormal findings in breast mammography. In Cardoso MJ, Arbel T, Carneiro G, et al., editors. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Spring International Publishing, Cham; 2017. p. 321– 329

[7] Becker AS, Marcon M, Ghafoor S, Wurnig MC, Frauenfelder T, and Boss A. Deep learning in mammography: Diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. *Investigate Radiology* 2016;52:434–440.

[8] Carneiro G, Nascimento J, and Bradley AP. Chapter 14 deep learning models for classifying mammogram exams containing unregistered multi-view images and segmentation maps of lesions. In Zhou SK, Greenspan H, and Shen D, editors, Deep

Learning for Medical Image Analysis. Academic Press; 2016. p. 321-339

[9] Kooi T and Karssemeijer N. Classifying symmetrical differences and temporal change in mammography using deep neural networks. *CoRR* 2017; abs/1703.07715.

[10] Mohamed AA, Berg WA, Peng H, Luo Y, Jankowitz RC, and Wu S. A deep learning method for classifying mammographic breast density categories. *Medical Physics* 2018;45(1):314–321.

[11] Ribli D, Horv´ath A, Unger Z, Pollner P, and Csabai I. Detecting and classifying lesions in mammograms with deep learning. *CoRR* 2017; abs/1707.08401.

[12] Shen L. End-to-end training for whole image breast cancer diagnosis using an all convolutional design. *CoRR* 2017, abs/1708.09427.

[13] Kyono T, Gilbert FJ, and van der Schaar M. MAMMO: A Deep Learning Solution for Facilitating Radiologist-Machine Collaboration in Breast Cancer Diagnosis. *ArXiv e-prints* 2018;abs/1811.02661.

[14] Gilbert F, Tucker L, Gillan MG, et al. The tommy trial: a comparison of tomosynthesis with mammography in the uknhs breast screening program. *Health Technology Assessment* 2015;19.

[15] Lakhani P, Prater AB, Hutson RK, et al. Machine learning in radiology: Applications beyond image interpretation. *Journal of the American College of Radiology* 2018;15(2):350 – 359.

[16] Syeda-Mahmood T. Role of big data and machine learning in diagnostic decision support in radiology. *Journal of the American College of Radiology* 2018;15(3, Part B):569 – 576.

[17] Lokate M, Stellato RK, Veldhuis WB, Peeters PHM, and van Gils CH. Age-related changes in mammographic density and breast cancer risk. *American Journal of Epidemiology* 2013;178(1):101–109.

[18] Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Annals of Internal Medicine*2003;138(3):168–175.

[19] Geras KJ, Wolfson S, Kim SG, Moy L, and Cho K. High resolution breast cancer screening with multi-view deep convolutional neural networks. *ArXiv e-prints* 2017;abs/1703.07047.

[20] Scutt D, Lancaster G, and Manning J. Breast asymmetry and predisposition to breast cancer. In *Breast cancer research* 2006;volume 8, page R14.

[21] Kendall A and Gal Y (2017). What uncertainties do we need in bayesian deep learning for computer vision? In Guyon I, Luxburg UV, Bengio S, et al., editors, *Advances in Neural Information Processing Systems 30*. Curran Associates, Inc.; 2017. p. 5574–5584.

[22] Wang G, Li W, Aertsen M, Deprest J, Ourselin S, and Vercauteren T. Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. *ArXiv e-Prints* 2018;abs/1807.07356.

[23] Sokolovskaya E, Shinde T, Ruchman R, et al. The effect of faster reporting speed for imaging studies on the number of misses and interpretation errors. *Journal of the American College of Radiology* 2015;12:683–688.

[24] Akselrod-Ballin A, Karlinsky L, Alpert S, Hasoul S, Ben-Ari R, and Barkan E. A region based convolutional network for tumor detection and classification in breast mammography. In Carneiro G, Mateus D, Peter L, et al., editors, *Deep Learning and Data Labeling for Medical Applications*. Cham: Springer International Publishing; 2016. p. 197-205.

[25] Jadoon M, Zhang Q, Ul Haq I, Butt S, and Jadoon A. Threeclass mammogram classification based on descriptive cnn features. *BioMed Research International* 2017;3640901.

[26] Jiao Z, Gao X, Wang Y, and Li J. A deep feature based framework for breast masses classification. *Neurocomputing* 2016; 197:221 – 231.

[27] Hepsag PU, Ozel SA, and Yazici A. Using deep learning for mammography classification. In *International Conference on Computer Science and Engineering (UBMK)* 2017;418–423.

[28] Platania R, Shams S, Yang S, Zhang J, Lee K, and Park SJ. Automated breast cancer diagnosis using deep learning and region of interest detection (bc-droid). In *Proceedings of the 8th*

ACM International Conference on Bioinformatics, Computational Biology and Health Informatics 2017;536–543.

[29] Samala R, Chan HP, Hadjiiski LM, Helvie AM, Cha K, and Richter CC. Multi-task transfer learning deep convolutional neural network: Application to computer-aided diagnosis of breast cancer on mammograms. In *Physics in Medicine and Biology* 2017; volume 62.

[30] Teare P, Fishman M, Benzaquen O, Toledano E, and Elnekave E. Malignancy Detection on Mammography Using Dual Deep Convolutional Neural Networks and Genetically Discovered False Color Input Enhancement. *Journal of Digital Imaging* 2017.

[31] Bredal IS, Kresen R, Skaane P, Engelstad KS, and Ekeberg I. Recall mammography and psychological distress. *European Journal of Cancer* 2013;49(4):805 – 811.

Figure Legends

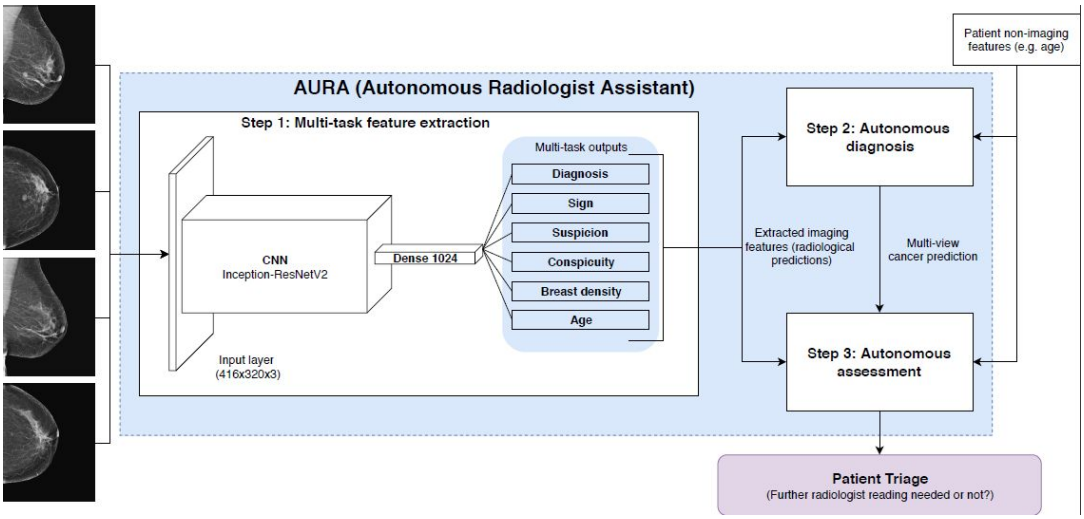

Fig 1. AURA system level illustration. In Step 1, all four of a patient's mammograms were passed to a multi-task feature extractor, i.e., a convolutional neural network (CNN) trained using multi-task learning (MTL) to emulate the mammographic predictions of a radiologist. In Step 2, the MTL predictions extracted from the individual mammograms were fused by a deep neural network to provide a multi-view diagnosis on the basis of both a patient's imaging and non-imaging features. Lastly, AURA considered the radiological predicted features (CNN multi-task outputs from Step 1), the multi-view cancer prediction (from Step 2), and the patient's

non-imaging features to issue a recommendation for *patient triage* to determine which patients could be autonomously diagnosed.

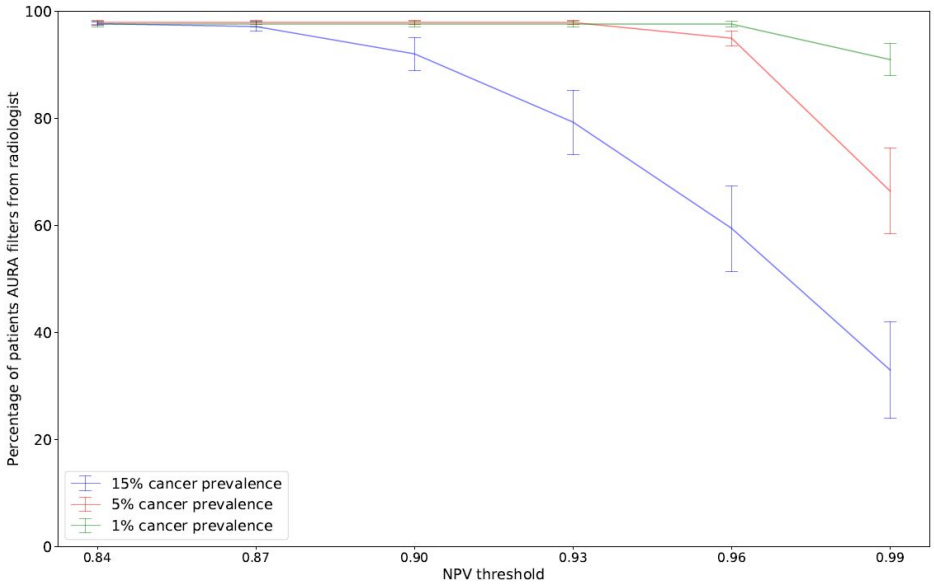

Fig 2. Percentage of patients that AURA filters from the radiologist across various NPV thresholds, as well as various cancer prevalence rates. 10-fold cross validation was used and 95% confidence intervals are shown.

Table 1: *Tommy* dataset patient distributions for breast density, age, mammographic signs, and mammographic conspicuity.

Attribute	**Value**	**Benign**	**Malignant**	**Prevalence**

All Patients		6164	1299	17.41%
Breast Density (along 10cm VAS)	0%- 25%	1789	454	20.24%
	25%- 50%	2942	554	15.85%
	50%- 75%	1075	231	17.69%
	75%- 100%	358	60	14.35%
Age	Under 50 years	625	51	7.54%
	50-60 years	3963	571	12.59%
	Over 60 years	1576	677	30.05%
Mammographic Signs	No sign of cancer	2688	76	2.75%
	Circumscribed mass	1469	119	7.49%
	Spiculated mass	62	434	87.50%
	Microcalcification	859	319	27.08%
	Distortion	144	145	50.17%
	Assymetrical Distortion	87	195	18.22%
Mammographic Conspicuity	None visible	2036	61	2.91%
	Barely visible	255	59	18.79%
	Visible, but not clear	1097	288	20.79%
	Clearly visible	2037	871	29.95%
Case Type	Assessment	4920	1201	19.62%
	Family History	1382	98	6.62%

Table 2: AURA Performance at the highest (15%) cancer prevalence. The odds ratio of receiving a negative diagnosis was calculated using logistic regression analysis. The AURA NPV was calculated as the joint NPV of the radiologist and the AURA classifier, where the average percentage filtered represents the number of patients that the AURA

classifier autonomously diagnosed (the rest received a diagnosis from a radiologist). Each value is shown with a 95% confidence interval.

Attribute	Value	Odds Ratio	AURA NPV	Avg. % Filtered
Breast Density (along 10cm VAS)	0%- 25%	159.21 ± 65.14	99.74% ± 0.15	31.83% ± 7.51
	25%- 50%	91.00 ± 35.28	99.59% ± 0.22	35.59% ± 6.16
	50%- 75%	58.36 ± 30.95	99.02% ± 0.28	32.09% ± 8.28
	75%- 100%	28.85 ± 21.16	98.18% ± 0.33	29.88% ± 9.39
Age	Under 50 years	83.50 ± 23.13	99.94% ± 0.04	41.82% ± 11.48
	50-60 years	69.84 ± 6.15	99.10% ± 0.40	36.14% ± 6.03
	Over 60 years	11.37 ± 0.68	92.02% ± 2.92	21.93% ± 5.68
Mammographic Signs	None visible	68.10 ± 8.82	99.85% ± 0.11	42.69% ± 8.11
	Circum. mass	40.79 ± 5.56	98.53% ± 0.69	37.11% ± 8.40
	Spic. mass	0.64 ± 0.05	34.29% ± 9.04	3.80% ± 2.04
	Microcalc.	7.07 ± 0.102	88.23% ± 3.82	28.03% ± 9.24
	Distortion	1.02 ± 0.09	49.77% ± 8.22	11.84% ± 8.47
	Assym. Distortion	65.88 ± 13.07	99.25% ± 0.26	29.33% ± 10.84
Mammographic Conspicuity	None visible	189.63 ± 43.52	99.93% ± 0.03	39.77% ± 10.76
	Barely visible	103.67 ± 50.31	99.04% ± 0.19	32.64% ± 9.35
	Visible, not clear	54.40 ± 8.33	98.21% ± 0.45	28.53% ± 8.45
	Clearly visible	57.16 ± 5.93	98.36% ± 0.51	27.55% ± 9.09

Case Type	Assessment	0.94 ± 0.02	48.23% ± 6.37	25.20% ± 4.12
	Family History	138.30 ± 33.31	99.42% ± 0.03	39.94% ± 11.70